ABSTRACT

Methods and systems are provided for partitioning data of a database or data store into several independent parts as part of a data mining process. The methods and systems use a mining application having content-based partitioning logic to partition the data. Once the data is partitioned, the partitioned data may be grouped and distributed to an associated processor for further processing. The mining application and content-based partitioning logic may be used in a computing system, including shared memory and distributed memory multi-processor computing systems. Other embodiments are described and claimed.